

What is claimed is:

1. A process for producing at least one haplotyped genome wide map, comprising the steps of:

- 5 (a) preparing chromosome maps associated with at least one chromosome; and
- (b) producing a portion of at least one haplotyped genome wide map based on the chromosome maps.

2. The process of claim 1, wherein the portion of at least one haplotyped genome wide map comprises at least one restriction site.

10 3. The process of claim 1, wherein less than all subparts of the haplotyped genome wide map are produced in step (b) as ordered or unordered sets of contigs.

4. The process of claim 1, wherein the chromosome maps are based on at least one single molecule map data set.

15 5. The process of claim 1, wherein the haplotyped genome wide map comprises two maps per chromosome associated with the at least one single molecule map data set.

20 6. The process of claim 4, wherein the at least one single molecule map data set has error rates as great as or smaller than: (i) about 10% error in distance between sites, (ii) about 20% missing sites, (iii) about 7% false sites and (iv) about 50% of sites closer than about 1 kB apart that are approximately indistinguishable.

25 7. The process of claim 4, wherein the at least one single molecule map data set consists of either Optical Mapping data or any single molecule ordered maps of polymorphic markers comprising at least one of restriction site polymorphisms, restriction length polymorphisms, insertions of bases, deletions of bases, and single nucleotide polymorphisms (SNPs).

8. The process of claim 4, wherein the at least one single molecule data set combines at least two single molecule map data sets comprising different restriction site markers combined into a single haplotyped genome wide map, wherein all restriction site markers are combined and wherein the restriction site markers are
5 distinguishable.

9. The process of claim 1, further comprising the step of determining a conditional probability density expression based on the chromosome maps associated with at least one chromosome.

10. The process of claim 9, wherein the probability density expression
10 is based on errors provided in at least one single molecule map data set.

11. The process of claim 1, further comprising the step of detecting substantially all site-based polymorphisms in the at least one haplotyped genome wide map.

12. The process of claim 1, further comprising the step of detecting substantially all interval-based polymorphisms in the at least one haplotyped genome wide map.
15

13. The process of claim 1, wherein steps (a) and (b) are performed within a particular time, and wherein the particular time is a sub-quadratic function of a number of sites associated with an input data.

20 14. A process for performing disease gene association based on at least one haplotyped genome wide map, comprising the steps of:

producing at least one haplotyped genome wide map per patient;
and

25 performing the disease gene association based on the produced at least one haplotyped genome wide map.

15. The process of claim 14, wherein the step of producing the at least one haplotyped genome wide map, comprises the sub steps of:

- (a) preparing chromosome maps associated with at least one chromosome; and
- (b) producing a portion of the at least one haplotyped genome wide map based on the chromosome maps.

5 16. The process of claim 14, wherein the portion of at least one haplotyped genome wide map comprises at least one restriction site.

17. The process of claim 14, wherein less than all subparts of the haplotyped genome wide map are produced in sub step (b) as ordered or unordered sets of contigs.

10 18. The process of claim 14, wherein the chromosome maps are based on at least one single molecule map data set.

19. The process of claim 14, wherein the haplotyped genome wide map comprises two maps per chromosome associated with the at least one single molecule map data set.

15 20. The process of claim 18, wherein the at least one single molecule map data set has error rates as great as or smaller than: (i) about 10% error in distance between sites, (ii) about 20% missing sites, (iii) about 7% false sites and (iv) about 50% of sites closer than about 1 kB apart that are approximately indistinguishable.

20 21. The process of claim 18, wherein the at least one single molecule map data set consists of either Optical Mapping data or any single molecule ordered maps of polymorphic markers comprising at least one of: restriction site polymorphism, restriction length polymorphisms, insertions of bases, deletions of bases, and single nucleotide polymorphisms (SNPs).

25 22. The process of claim 18, wherein the at least one single molecule data set combines at least two single molecule map data sets comprising different restriction site markers is assembled into a single haplotyped genome wide map, wherein

all restriction site markers are combined and wherein the restriction site makers are distinguishable.

23. The process of claim 14, further comprising the step of determining a conditional probability density expression based on the chromosome maps associated with at least one chromosome.

24. The process of claim 23, wherein the probability density expression is based on errors provided in at least one single molecule map data set.

25. The process of claim 14, further comprising the step of detecting substantially all site-based polymorphisms in the at least one haplotyped genome wide map.

26. The process of claim 14, further comprising the step of detecting substantially all interval-based polymorphisms in the at least one haplotyped genome wide map.

27. The process of claim 14, wherein steps (a) and (b) are performed within a particular time, and wherein the particular time is a sub-quadratic function of a number of sites associated with an input data.

28. A process for producing at least one genotyped genome wide map comprising the steps of:

- (a) preparing chromosome maps associated with at least one chromosome; and
- (b) producing a portion of at least one genotyped genome wide map based on the chromosome maps.

29. The process of claim 28, wherein the portion of at least one genotyped genome wide map comprises at least one restriction site.

30. The process of claim 28, wherein less than all subparts of the genotyped genome wide map are produced in step (b) as ordered or unordered sets of contigs.

31. The process of claim 28, wherein the chromosome maps are based
5 on at least one single molecule map data set.

32. The process of claim 28, wherein the genotyped genome wide map comprises two maps per chromosome is assembled from at least one single molecule map data set.

33. The process of claim 31, wherein the at least one single molecule
10 map data set has error rates as great as or smaller than: (i) about 10% error in distance between sites, (ii) about 20% missing sites, (iii) about 7% false sites and (iv) about 50% of sites closer than about 1 kB apart that are approximately indistinguishable.

34. The process of claim 31, wherein the at least one single molecule map data set consists of either Optical Mapping data or any single molecule ordered maps
15 of polymorphic markers comprising at least one of restriction site polymorphisms, restriction length polymorphisms, insertions of bases, deletions of bases, and single nucleotide polymorphisms (SNPs).

35. The process of claim 31, wherein the at least one single molecule data set combines at least two single molecule map data sets comprising different
20 restriction site markers is assembled into a single genotyped genome wide map, wherein all restriction site markers are combined and wherein the different restriction site markers are distinguishable.

36. The process of claim 28, further comprising the step of determining a conditional probability density expression based on the chromosome maps
25 associated with at least one chromosome.

37. The process of claim 36, wherein the probability density expression is based on errors provided in at least one single molecule map data set.

38. The process of claim 28, further comprising the step of detecting substantially all site-based polymorphisms in the at least one genotyped genome wide map.

39. The process of claim 28, further comprising the step of detecting substantially all interval-based polymorphisms in the at least one genotyped genome wide map.

40. The process of claim 28, wherein steps (a) and (b) are performed within a particular time limit, and wherein the particular time is a sub-quadratic function of a number of sites associated with an input data.

41. A software arrangement which, when executed on a processing device, configures the processing device to produce at least one haplotyped genome wide map comprising the steps of:

(a) preparing chromosome maps associated with at least one chromosome; and

(b) producing a portion of at least one haplotyped genome wide map based on the chromosome maps.

42. The software arrangement according to claim 41, wherein the portion of at least one haplotyped genome wide map comprises at least one restriction site.

43. The software arrangement according to claim 41, wherein less than all subparts of the haplotyped genome wide map are produced in step (b) as ordered or unordered sets of contigs.

44. The software arrangement according to claim 41, wherein the chromosome maps are based on at least one single molecule map data set.

45. The software arrangement according to claim 41, wherein the haplotyped genome wide map comprises two maps per chromosome is assembled from the at least one single molecule map data set

46. The software arrangement according to claim 44, wherein the at least one single molecule map data set has error rates as great as or smaller than: about 10% error in distance between sites, about 20% missing sites, about 7% false sites and about 50% of sites closer than about 1 kB apart that are indistinguishable.

5 47. The software arrangement according to claim 44, wherein the at least one single molecule map data set consists of either Optical Mapping data or any single molecule ordered maps of polymorphic markers comprising at least one of restriction site polymorphisms, restriction length polymorphisms, insertions of bases, deletions of bases, single nucleotide polymorphisms (SNPs).

10 48. The software arrangement according to claim 44, wherein the at least one single molecule map data sets comprising different restriction site markers are assembled into a single haplotyped genome wide map wherein all restriction site markers are combined and wherein the restriction site markers can be distinguished.

15 49. The software arrangement according to claim 41, further comprising determining a conditional probability density expression.

50. The software arrangement according to claim 49, wherein the probability density expression is based on errors provided in at least one single molecule map data set.

20 51. The software arrangement according to claim 41, wherein substantially all site based polymorphisms are detected in the at least one haplotyped genome wide map.

52. The software arrangement according to claim 41, wherein substantially all interval-based polymorphisms are detected in the at least one haplotyped genome wide map.

25 53. The software arrangement according to claim 41, wherein steps (a) and (b) are performed within a particular time limit, and the particular time is a sub-quadratic function of a number of sites associated with an input data.

54. The software arrangement according to claim 41, further comprising performing a disease gene association study based on at least one haplotyped genome wide map per patient.

55. A software arrangement which, when executed on a processing device, configures the processing device to perform disease gene association based on at least one haplotyped genome wide map per patient, comprising the steps of:

- (a) producing at least one haplotyped genome wide map per patient ; and
- (b) performing the disease gene association based on the produced at least one haplotyped genome wide map.

56. The software arrangement according to claim 55, wherein the step of producing the at least one haplotyped genome wide map, comprises the sub steps of:

- (a) preparing chromosome maps associated with at least one chromosome; and
- (b) producing a portion of the at least one haplotyped genome wide map based on the chromosome maps.

57. The software arrangement according to claim 56, wherein the portion of at least one haplotyped genome wide map comprises at least one restriction site.

58. The software arrangement according to claim 56, wherein less than all subparts of the haplotyped genome wide map are produced in sub step (b) as ordered or unordered sets of contigs.

59. The software arrangement according to claim 56, wherein the chromosome maps are based on at least one single molecule map data set.

60. The software arrangement according to claim 56, wherein the haplotyped genome wide map comprises two maps per chromosome is assembled from at least one single molecule map data set.

61. The software arrangement according to claim 59, wherein the at least one single molecule map data set has error rates at large as or smaller than: (i) about 10% error in distance between sites, (ii) about 20% missing sites, (iii) about 7% false sites and (iv) about 50% of sites closer than about 1 kB apart that are approximately 5 indistinguishable.

62. The software arrangement according to claim 59, wherein the at least one single molecule map data set consists of either Optical Mapping data or any single molecule ordered maps of polymorphic markers comprising at least one of restriction site polymorphisms, restriction length polymorphisms, insertions of bases, 10 deletions of bases, and single nucleotide polymorphisms (SNPs).

63. The software arrangement according to claim 59, wherein the at least one single molecule data set combines at least two single molecule map data sets comprising different restriction site markers is assembled into a single haplotyped genome wide map, wherein all restriction site markers are combined and wherein the 15 different restriction site markers are distinguishable.

64. The software arrangement according to claim 56, further comprising the step of determining a conditional probability density expression based on the chromosome maps associated with at least one chromosome.

65. The software arrangement according to claim 64, wherein the 20 probability density expression is based on errors provided in at least one single molecule map data set.

66. The software arrangement according to claim 56, further comprising the step of detecting substantially all site-based polymorphisms in the at least one haplotyped genome wide map.

67. The software arrangement according to claim 56, further comprising the step of detecting substantially all interval-based polymorphisms in the at 25 least one haplotyped genome wide map.

68. The software arrangement according to claim 56, wherein steps (a) and (b) are performed within a particular time limit, and wherein the particular time is a sub-quadratic function of a number of sites associated with an input data.

69. A software arrangement which, when executed on a processing device, configures the processing device to produce at least one genotyped genome wide map comprising the steps of:

- (a) preparing chromosome maps associated with at least one chromosome; and
- (b) producing a portion of at least one genotyped genome wide map based on the chromosome maps.

70. The software arrangement according to claim 69, wherein the portion of at least one genotyped genome wide map comprises at least one restriction site.

71. The software arrangement according to claim 69, wherein less than all subparts of the genotyped genome wide map are produced in step (b) as ordered or unordered sets of contigs.

72. The software arrangement according to claim 69, wherein the chromosome maps are based on at least one single molecule map data set.

73. The software arrangement according to claim 69, wherein the genotyped genome wide map comprises two maps per chromosome associated with the at least one single molecule map data set

74. The software arrangement according to claim 72, wherein the at least one single molecule map data set has error rates as large as or smaller than: about 10% error in distance between sites, about 20% missing sites, about 7% false sites and about 50% of sites closer than about 1 kB apart that are indistinguishable.

75. The software arrangement according to claim 72, wherein the at least one single molecule map data set comprises Optical Mapping data or any single molecule ordered maps of polymorphic markers comprising any of restriction site

polymorphisms, restriction length polymorphisms, insertions of bases, deletions of bases, single nucleotide polymorphisms (SNPs).

76. The software arrangement according to claim 72, wherein at least one single molecule map data set comprising different restriction site markers are assembled into a single genotyped genome wide map wherein all restriction site markers are combined and wherein the different restriction site markers can be distinguished.

5 77. The software arrangement according to claim 69, further comprising determining a conditional probability density expression.

10 78. The software arrangement according to claim 77, wherein the probability density expression is based on errors provided in at least one single molecule map data set.

79. The software arrangement according to claim 69, wherein substantially all site based polymorphisms are detected in the at least one genotyped genome wide map.

15 80. The software arrangement according to claim 69, wherein substantially all interval-based polymorphisms are detected in the at least one genotyped genome wide map.

20 81. The software arrangement according to claim 69, wherein steps (a) and (b) are performed within a particular time limit, and the particular time is a sub-quadratic function of a number of sites associated with an input data.

82. The software arrangement according to claim 69, further comprising performing a disease gene association study based on at least one genotyped genome wide map per patient.

25 83. A system for producing at least one haplotyped genome wide map comprising a storage medium wherein the storage medium includes software that is executed to perform the steps of:

- (a) preparing chromosome maps associated with at least one chromosome; and
- (b) producing a portion of at least one haplotyped genome wide map based on the chromosome maps.

- 5 84. The system of claim 83, wherein the portion of at least one haplotyped genome wide map comprises at least one restriction site.
- 10 85. The system of claim 83, wherein less than all subparts of the haplotyped genome wide map are produced in step (b) as ordered or unordered sets of contigs.
- 15 86. The system of claim 83, wherein the chromosome maps are based on at least one single molecule map data set.
- 20 87. The system of claim 83, wherein the haplotyped genome wide map comprises two maps per chromosome is assembled from the at least one single molecule map data set.
- 25 88. The system of claim 86, wherein the at least one single molecule map data set has error rates as large as or smaller than: (i) about 10% error in distance between sites, (ii) about 20% missing sites, (iii) about 7% false sites and (iv) about 50% of sites closer than about 1 kB apart that are approximately indistinguishable.
- 30 89. The system of claim 86, wherein the at least one single molecule map data set comprises Optical Mapping data or any single molecule ordered maps of polymorphic markers comprising at least one of restriction site polymorphisms, restriction length polymorphisms, insertions of bases, deletions of bases, and single nucleotide polymorphisms (SNPs).
- 35 90. The system of claim 86, wherein the at least one single molecule data set combines at least two single molecule map data sets comprising of different restriction site markers is assembled into a single haplotyped genome wide map, wherein

all restriction site markers are combined and wherein the different restriction site markers are distinguishable.

91. The system of claim 83, further comprising the step of determining a conditional probability density expression based on the chromosome maps associated with at least one chromosome.

92. The system of claim 91, wherein the probability density expression is based on errors provided in at least one single molecule map data set.

93. The system of claim 83, further comprising the step of detecting substantially all site-based polymorphisms in the at least one haplotyped genome wide map.

94. The system of claim 83, further comprising the step of detecting substantially all interval-based polymorphisms in the at least one haplotyped genome wide map.

95. The system of claim 83, wherein steps (a) and (b) are performed within a particular time limit, and wherein the particular time is a sub-quadratic function of a number of sites associated with an input data.

96. A system for performing disease gene association based on at least one haplotyped genome wide map per patient, comprising a storage medium wherein the storage medium includes software that is executed to perform the steps of:

- 20 (a) producing at least one haplotyped genome wide map per patient; and
(b) performing the disease gene association based on the produced at least one haplotyped genome wide map.

97. The system of claim 96, wherein the step of producing the at least 25 one haplotyped genome wide map, comprises the sub steps of:

- (a) preparing chromosome maps associated with at least one chromosome; and
- (b) producing a portion of the at least one haplotyped genome wide map based on the chromosome maps.

5 98. The system of claim 97, wherein the portion of at least one haplotyped genome wide map comprises at least one restriction site.

99. The system of claim 97, wherein less than all subparts of the haplotyped genome wide map are produced in sub step (b) as ordered or unordered sets of contigs.

10 100. The system of claim 97, wherein the chromosome maps are based on at least one single molecule map data set.

101. The system of claim 97, wherein the haplotyped genome wide map comprises two maps per chromosome is assembled from at least one single molecule map data set.

15 102. The system of claim 100, wherein the at least one single molecule map data set has error rates as large as or smaller than: (i) about 10% error in distance between sites, (ii) about 20% missing sites, (iii) about 7% false sites and (iv) about 50% of sites closer than about 1 kB apart that are approximately indistinguishable.

20 103. The system of claim 100, wherein the at least one single molecule map data set comprises Optical Mapping data or any single molecule ordered maps of polymorphic markers comprising at least one of restriction site polymorphisms, restriction length polymorphisms, insertions of bases, deletions of bases, and single nucleotide polymorphisms (SNPs).

25 104. The system of claim 100, wherein the at least one single molecule data set combines at least two single molecule map data sets comprising different restriction site markers is assembled into a single haplotyped genome wide map, wherein

~~all restriction site markers are combined and wherein the different restriction site markers are distinguishable.~~

105. The system of claim 97, further comprising the step of determining a conditional probability density expression based on the chromosome maps associated 5 with at least one chromosome.

106. The system of claim 105, wherein the probability density expression is based on errors provided in at least one single molecule map data set.

107. The system of claim 97, further comprising the step of detecting substantially all site-based polymorphisms in the at least one haplotyped genome wide 10 map.

108. The system of claim 97, further comprising the step of detecting substantially all interval-based polymorphisms in the at least one haplotyped genome wide map.

109. The system of claim 97, wherein steps (a) and (b) are performed 15 within a particular time limit, and wherein the particular time is a sub-quadratic function of a number of sites associated with an input data.

110. A system for producing at least one genotyped genome wide map comprising a storage medium wherein the storage medium includes software that is executed to perform the steps of:

- 20 (a) preparing chromosome maps associated with at least one chromosome; and
(b) producing a portion of at least one genotyped genome wide map based on the chromosome maps.

111. The system of claim 110, wherein the portion of at least one 25 genotyped genome wide map comprises at least one restriction site.

112. The system of claim 110, wherein less than all subparts of the genotyped genome wide map are produced in step (b) as ordered or unordered sets of contigs.

113. The system of claim 110, wherein the chromosome maps are based
5 on at least one single molecule map data set.

114. The system of claim 110, wherein the genotyped genome wide map comprises two maps per chromosome assembled from at least one single molecule map data set.

115. The system of claim 113, wherein the at least one single molecule
10 map data set has error rates as large as or smaller than: (i) about 10% error in distance between sites, (ii) about 20% missing sites, (iii) about 7% false sites and (iv) about 50% of sites closer than about 1 kB apart that are approximately indistinguishable.

116. The system of claim 113, wherein the at least one single molecule map data set comprises Optical Mapping data or any single molecule ordered maps of
15 polymorphic markers comprising at least one of restriction sites, restriction length polymorphisms, insertions of bases, deletions of bases, and single nucleotide polymorphisms (SNPs).

117. The system of claim 113, wherein the at least one single molecule data set combines at least two single molecule map data sets comprising different
20 restriction site markers is assembled into a single genotyped genome wide map, wherein all restriction site markers are combined and wherein the different restriction site markers are distinguishable.

118. The system of claim 110, further comprising the step of determining a conditional probability density expression based on the chromosome maps
25 associated with at least one chromosome.

119. The system of claim 118, wherein the probability density expression is based on errors provided in at least one single molecule map data set.

120. The system of claim 110, further comprising the step of detecting substantially all site-based polymorphisms in the at least one genotyped genome wide map.

121. The system of claim 110, further comprising the step of detecting substantially all interval-based polymorphisms in the at least one genotyped genome wide map.

122. The system of claim 110, wherein steps (a) and (b) are performed within a particular time limit, and wherein the particular time is a sub-quadratic function of a number of sites associated with an input data.